

Original citation:

Fuller, Steve (2018) *The brain as artificial intelligence : prospecting the frontiers of neuroscience*. AI & SOCIETY .doi:[10.1007/s00146-018-0820-1](https://doi.org/10.1007/s00146-018-0820-1)

Permanent WRAP URL:

<http://wrap.warwick.ac.uk/101005>

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work of researchers of the University of Warwick available open access under the following conditions.

This article is made available under the Creative Commons Attribution 4.0 International license (CC BY 4.0) and may be reused according to the conditions of the license. For more details see: <http://creativecommons.org/licenses/by/4.0/>

A note on versions:

The version presented in WRAP is the published version, or, version of record, and may be cited as it appears here.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk



The brain as artificial intelligence: prospecting the frontiers of neuroscience

Steve Fuller¹

Received: 12 March 2017 / Accepted: 24 January 2018
© The Author(s) 2018. This article is an open access publication

Abstract

This article explores the proposition that the brain, normally seen as an organ of the human body, should be understood as a biologically based form of artificial intelligence, in the course of which the case is made for a new kind of ‘brain exceptionalism’. After noting that such a view was generally assumed by the founders of AI in the 1950s, the argument proceeds by drawing on the distinction between science—in this case neuroscience—adopting a ‘telescopic’ or a ‘microscopic’ orientation to reality, depending on how it regards its characteristic investigative technologies. The paper concludes by recommending a ‘microscopic’ yet non-reductionist research agenda for neuroscience, in which the brain is seen as an underutilised organ whose energy efficiency is likely to outstrip that of the most powerful supercomputers for the foreseeable future.

Keywords Brain exceptionalism · Cognitive agriculture · Microscope · Neural nets · Neuroscience · Reductionism · Supercomputer · Telescope

In what follows, I explore the proposition that the brain, normally seen as an organ of the human body, should be understood as a biologically based form of artificial intelligence (AI). As I observe in the “Introduction”, this proposition was assumed by the founders of AI in the 1950s, though it has been generally side-lined over the course of AI’s history. However, advances in both neuroscience and more conventional AI make it interesting to consider the issue anew. The main body of the paper approaches the matter from a distinction that the philosopher Bas van Fraassen has drawn in terms of science adopting a ‘telescopic’ or a ‘microscopic’ orientation to reality. The history of neuroscience has exhibited both tendencies from its inception, not least in terms of the alternative functions performed by the field’s characteristic technologies. Appreciating the full implications of this distinction requires escaping from the ‘reductionist’ problematic that continues to haunt philosophical discussions of neuroscience’s aspirations as a mode of inquiry. As becomes clear by the Conclusion, my own preference is for an ambitious ‘microscopic’ agenda for neuroscience, which in the

long term may see organically grown neural networks—if not full-fledged brains—carrying out many if not most of the functions that nowadays are taken to be the purview of silicon-based computers. In this respect, I am arguing for a new kind of ‘brain exceptionalism’, one based not on the brain’s natural mysteries but on its relative energy efficiency vis-à-vis competing (silicon) technologies.

1 Introduction: the brain as a work in progress

The history of artificial intelligence (AI) is normally told in terms of three research strategies that coalesced in the wake of the Macy Conferences, which were first held at Dartmouth College in 1956: one focused on neural networks, one on cybernetic systems and one simply on computational power (Gardner 1985: Chap. 6). The third, which carried the least ontological baggage, became the dominant understanding of AI. A striking feature of all three strategies—which perhaps helps to explain the dominance of the third—is the failure to clearly distinguish *modelling* and *making* the target phenomenon, ‘intelligence’ (Dupuy 2000: Chap. 2). This may have been to do with the preponderance of mathematicians among the AI founders, who privileged the search for analytic clarity over the need to provide concrete instantiation,

✉ Steve Fuller
s.w.fuller@warwick.ac.uk

¹ Auguste Comte Chair in Social Epistemology, Department of Sociology, University of Warwick, Coventry CV4 7AL, UK

sometimes to the exasperation of interloping physicists (Malapi-Nelson 2017: Chaps. 7–8).

To be sure, this persistent ambiguity has contributed to the Turing Test's iconic status as popular culture's portal into the world of AI. After all, the test's fascination rests on the difficulty that humans normally have in distinguishing a 'natural' from an 'artificial' intelligence (i.e. a fellow human from an android) simply based on the candidate being's performance in response to their questions. It led one of the AI pioneers, psychiatrist Ross Ashby, to quip that passing the Turing test simply meant that questioner and respondent had established a common standard of something they agree to call 'intelligence' without saying what it is. For Ashby, the 'brain' was whatever could reliably enable an entity to pass such a test, regardless of its material composition (Malapi-Nelson 2017: Chap. 6). This orientation to intelligence is rather like the economist's view that a good's value is simply the price that it fetches in a free market or, perhaps more to the point, the Popperian view that what both scientists and lay people mean by 'true' is simply whatever passes critical scrutiny. In each of these cases, the normative standard—intelligence, value and truth—functions as a reversible convention rather than as a fixed essence.

However, the blurring of the natural/artificial distinction cannot simply be reduced to the abstractness—or ontological neutrality—of the AI research agenda itself. After all, the one arguably 'natural' focus of AI research, the brain, has itself always been regarded within this tradition a work in progress, an artefact in the making. Indeed, Ashby published a book in 1952 with the bold title, *Design for a Brain*. Moreover, the original AI researchers specifically focused on the brain—Warren McCulloch and Walter Pitts—did not try to model the entire organ in terms of its multiple known functionalities, which one might think would have been the most logical way to go and is probably the dominant sense in which AI researchers today think about the prospect of computers possessing 'brains'. Instead they constructed an idealised version of the biological brain's functional equivalent to an atom, the neuron. Vast combinations of these neurons were portrayed as engaged in the parallel processing of data from a variety of sensory sources, resulting in emergent patterns of co-activity which over time become so integrated that they then—and only then—constitute a 'brain'.

In effect, McCulloch and Pitts had proposed to grow a brain in a silicon setting. The project was initially met with considerable scepticism by early AI's most constructive critic, John von Neumann, but was later revived with much enhanced computer power in the 1980s under the rubrics of 'connectionism' and 'parallel distributed processing'. It continues to enjoy support today, though still far from meeting its initial promise (Malapi-Nelson 2017: Chap. 7; Boden 2006: Chap. 12). What remains striking about this approach—and animates the spirit of my argument—is its

assumption that the way in which biological brains instantiate intelligence captures something deep about intelligence itself, which remains untapped if we only focus on the brain's known functionalities, as these may simply reflect the specific pathways through which brains have so far evolved.

In other words, what McCulloch and Pitts had captured was the idea that the organic is *organized*—not only that brains perform as they do because of the environments in which they have been so far placed, but also that in different environments brains would behave in (potentially better) ways precisely because of their material composition. The latter point, which also informs the thinking behind contemporary 'synthetic biology', envisages the genes and cells that compose 'natural' living beings as literal building blocks—self-contained energy modules, if you will—that could be 'organized' to produce new (and improved) bio-based edifices by drawing on the expertise of both biologists and engineers (Church and Regis 2012; cf.; Fuller 2016a). The most obvious precedent for this line of thought is capitalism's default understanding of labour as an underutilised productive force or unexploited potential that may be improved by 'better organisation'. Indeed, one of the British AI pioneers, Stafford Beer, picked up on this point to become one of the early exemplars of the 'management guru' (Pickering 2010: Chap. 6). The final section of this paper revisits this sensibility via 'cognitive agriculture'.

To be sure, more than a half-century after AI's original encounters with the brain, biochemistry continues to offer much the same appraisal of the organ's self-creative capacity, even in its natural carbon-centred terms: human genes produce 100 billion neurons that are only locally organised at an individual's birth but become more globally integrated through repeated experience and feedback over the individual's lifetime (Williams and Frausto da Silva 2007: 483). Observations of this sort have helped to fuel the recent interest in 'epigenetics', with its promise of training up an always only partially formed brain via a cocktail of chemicals and external stimuli, even to the point of allowing for an 'adult neurogenesis' (Rubin 2009).

But more directly relevant for our purposes, this general 'neuroplastic' understanding of the brain casts doubt on the popular but crude distinction that the brain is 'hardware' to the mind's 'software'. The metaphor jars because it suggests too strong an association between computer hardware and the 'blank slate' conception of the mind traditionally attributed to Aristotle and Locke. In recent times, this view has been vilified by the evolutionary psychologist Steven Pinker (2002) for leaving the impression that the brain's powers are reducible to the sum of the algorithms that are programmed into it once it comes into the world. To be sure, the ease with which we speak of 'erasing' data from a computer drive points to such a slate-like conceptualisation of hardware. Thus, those who like Ray Kurzweil nowadays fixate

on Moore's law—which points to an exponential growth in silicon-based computational power—may be unwittingly exemplifying Pinker's point. Here computer hardware is literally little more than a platform for software, the latter taken to be ultimately driving the AI project, which will get easier over time as Moore's law plays itself out and the platform becomes increasingly capable of doing more with less.

Pinker's own argument—perhaps a bit too influenced by genetics—emphasises the brain's capacity to *resist* certain attempts at the sort of customisation that he associates with utopian political sensibilities, starting with Rousseau and including Marx and even B.F. Skinner, whose own algorithmically ordered utopia was governed by 'schedules of reinforcement' (Fuller 2006: 172–173, 196–201). However, the argument can be flipped to imply that the brain's capacity, while not inherently 'blank', has yet to be tapped in the right ways. In other words, utopians of yore may have failed simply because they did not (know how to) intervene in brain processes in ways that would enable their policies to stick in their recipients' minds as bases for action. But it does not follow that human brains are incapable of realising something similar in scale and scope to the utopians' blueprints. In that case, a more appropriate contrast for understanding the brain-mind relationship may be the *wetware/dryware* distinction, which is drawn by nanotechnologists to distinguish the organic base ('wetware') from the prosthetic attachments ('dryware'), which when taken together turn an organism into a cyborg.

The advantage of this distinction is that it does not presuppose a clean ontological split between brain and mind. Indeed, the original proposal for a 'philosophy of technology' in the late nineteenth century by Ernst Kapp assumed that over the course of human evolution, the powers of our brains are being enhanced by the multiplication and amplification of our senses, which may eventually involve their wetware components being replaced by dryware extensions (Brey 2000). The neuroscientist David Eagleman has updated and concretised this idea by suggesting that in the not too distant future clothing may be designed to interface directly with the brain, thereby providing portals for their wearers to process additional sensory data on a regular basis (Mason 2015). This prospect should be kept in mind as we consider what alternative approaches to neuroscience say about the sort of entity that we take the brain to be.

2 Two technological approaches to the brain: the telescope and the microscope

Thomas Kuhn (1962) famously likened a scientific revolution to a Gestalt switch. What he meant was that often it takes only a slight shift in perspective to cause a radical

shift in understanding. In principle, no new facts are needed, just a new sense of the logic of their arrangement. At the moment, the research aspirations of neuroscience—at least as popularly represented—often seem captive to a philosophical impasse, the problematic of *reductionism*, which requires a Gestalt switch.

There is a real disagreement between those who envisage neuroscience as potentially expanding our conception of reality and those who see it as an auxiliary science, ultimately telling us little that we did not already know, except when brains do not enable their bearers to do what is expected of them. On the one hand, those who argue for an epistemically strong neuroscience agenda often suggest that we are simply our brains; on the other, those who resist neuroscience's delusions of grandeur counter that we simply use our brains to be who we are, which in some important sense always escapes the confines of the brain. Philosophers tend to stereotype this disagreement as 'materialism versus dualism', the former functioning as the unofficial ideology of science and the latter of religion. Put more precisely, the 'materialists' on this view uphold a universal scientism, whereas the 'dualists' believe that science must ultimately yield to a more spiritual way of seeing the world. Thus, the former are dubbed 'reductionists' and the latter 'anti-reductionists'.

But as with so many other philosophical dichotomies, the scholastic familiarity of this one should breed contempt. My antidote is to turn the materialist-dualist polarity 90° on its axis, which would replace the reductionism problematic with a more explicitly sociological one about the sort of world that we would wish to inhabit, a problematic in which the brain continues to play a central role. In what follows, I characterise the two new poles in terms of the axial rotation that they require.

When 'materialists' sound like they are claiming that we are *no more* than our brains, they should be heard as meaning that we are *no less*. In other words, materialism in neuroscience would be better understood not as reducing the complex person to the activity of single organ, but as highlighting that organ's distinctive capacity to enable its bearer to think and do a much wider range of things than s/he normally does. This was the spirit in which Hilary Putnam (1982) originally proposed the 'brains in a vat' scenario as a stronger version of the Cartesian problem of scepticism. Whereas Descartes imagined an evil demon capable of fabricating the world we actually experience, Putnam's brains (*sans* demon) could do much more than that, precisely because they were hooked up to a machine that enabled the brains to perceive whatever *they* thought as reality. The intriguing suggestion here is that what we normally regard as the 'external' character of reality—be it caused by nature, society or some demon—may simply reflect limits placed on the full exploitation of the brain's powers. This way of looking at things would stress the extent to which the brain

is held back by surroundings that fail to stimulate it in some relevant ways. In this context, proposals concerning ‘smart environments’ and ‘social enhancement’, be they made by B.F. Skinner, Donald Norman or contemporary transhumanists can be seen as complementing the earlier discussion of David Eagleman’s work on wearable sensory technologies (cf. Cabrera 2015).

Similarly, those ‘dualists’ who stress that we are not reducible to our brains are not really positing a spiritual ‘I’ that oversees the brain’s activities and literally uses the organ for its own ends. Rather they should be interpreted as meaning that even before we do any neuroscience, we already know what our brains are supposed to do—albeit by indirect means, such as observed behaviour or introspective states. This evidence is provided by the brain’s constant interaction with the rest of the body in which it is located, and that body with the larger environment, which includes other beings with brains. In that respect, the brain might be regarded as a synecdoche for the mind—that is, part of an ensemble of entities, which are orchestrated to bring about thinking and thoughtful action.

Clark (2008) and Bennett and Hacker (2013) offer complementary glosses on this position, the former presented as a positive research programme in cognitive science and the latter as a systematic critique of contemporary cognitive neuroscience. In both cases, the phrase ‘extended mind’ is not inappropriate, since the idea is to shift the parameters of the mind from the brain itself to the other entities with which a brain co-produces a meaningful reality. In that case, neuroscience should be about explaining the specific structures and processes in the brain and nervous system that contribute to those publicly recognisable states of being ‘mindful’. Such an approach would stress the extent to which a brain enables its bearer to function in its normal environment. How that actually happens with regard to the neurophysiology of an individual brain is bound to vary, perhaps even significantly, depending on personal history. When the brain-bearer is behaving normally, those differences probably do not matter, but abnormal public behaviour may require intensive brain-based investigation.

How does this refocusing of materialism and dualism constitute a ‘90°’ turn? The underlying idea is that each side of divide retains a key aspect of its position but also adopts a key aspect of the other position. Thus, the materialist is still fixed on the brain as the locus of the mind, and the dualist on the mind as something that escapes the brain. That’s the part of the old positions that remains intact. However, whereas older materialists would have dismissed out of hand the existence of a paranormal realm for its lack of conformity to normal brain function, my revisionist view of materialism is open to locating the paranormal realm in areas of the brain that remain ‘unexplored’, in the sense of not having received the relevant sensory contact. This

involves a significant concession to classical dualism, but it is perhaps not so different from what passed for *monism* as a scientific philosophy in the late nineteenth and early twentieth centuries, which included such hybrid metaphysical positions as ‘panpsychism’ (Fechner), ‘hylozoism’ (Haeckel) and ‘energeticism’ (Ostwald), all of which were open—to varying degrees—to probing the frontiers of consciousness (Weir 2012). Similarly, whereas older dualists would have dismissed out of hand any spatiotemporal specification of the mental and perhaps even spiritual properties that are by definition unconfined by brain processes, this revisionist dualism approximates a Neo-Heideggerian or Neo-Wittgensteinian view of the mind—familiar, say, from the sociology of the scientific knowledge—as a distributed entity that is defined by people and things located in specifiable regions of space–time which stand in certain mutually recognisable relations to each other, only some of which are explicitly codified but the violation of which can be verified by any of the relevant parties to those relations.

I said earlier that as a result of this 90° turn, the reductionist problematic in neuroscience morphs into a more general question of the sort of world in which we would wish to live. The ‘new look’ materialist adopts what I have called, in another context, a *proactionary* attitude towards the brain, whereas the ‘new look’ dualist adopts a *precautionary* attitude towards the world (Fuller and Lipinska 2014). The former tends to see the brain as normally underutilised and hence always in need of ‘enhancement’ or ‘improvement’ to realise its full potential, while the latter sees the brain as normally utilised just as it should but there remains an open question as to which other entities in the brain’s environment help to sustain this normal functionality. A good way to capture the difference in bio-evolutionary terms is to see the proactionaries as stressing the fact that in the roughly 40,000 years since the human brain acquired its current organic form, humanity’s orientation to the world has changed radically, ever more rapidly as we get closer to the present. This phenomenon is not unreasonably seen from a historical standpoint as the product of a direct and indirect re-purposing of the brain (Smail 2008). In contrast, the precautionaries place greater emphasis on the relatively long duration of this process, as well as the increasing amount of ecological disruption that it has caused in recent times, as humanity has become more insistent on expediting the process by turning its cerebral emissions (aka ‘ideas’) into reality, regardless of the long-term consequences for the other entities on which the brain continues to depend.

But let us step back from politics and think about this fork in the road in the future of neuroscience in terms of the style of inquiry that each side implies. Thirty-five years ago, the philosopher of science Bas van Fraassen (1981) distinguished the *telescope* and the *microscope* as instruments for inquiring into the nature of things. The distinction suggested

two modes of epistemic enhancement through technology. The telescope magnifies entities that we can already see to some extent, the suggestion being that the instrument merely fills in the details of our understanding of an object that we already know—however indistinctly—with the naked eye, such as a planet or a star. It enhances our knowledge without challenging our conceptual framework. In contrast, the microscope provides us access to entities, such as germs or atoms, to which we not only lacked prior access but also might have previously regarded as figments of the imagination. In this way, microscopic discoveries may cause us to rethink our conceptual framework.

This was quite a novel way of distinguishing the two instruments. After all, *prima facie* the difference between the telescope and the microscope is that the former makes far away big objects visible while the latter makes nearby small objects visible. In other words, both instruments are ordinarily seen as aiming to present their objects on a common plane of visibility, which is what normally passes for empirical reality. However, in van Fraassen's reformulation, the two instruments differ substantially in ontological import: the microscope as a technology compels a reorientation in world-view that the telescope does not. Current foundational debates about the scope of neuroscience pivot between van Fraassen's take on these two instruments. On the one hand, our inquiry may be driven by what it is about the brain (and nervous system) that enables us to do the sorts of things and think the sorts of thoughts that we can already do—as well as what inhibits or distorts them. That is a 'telescopic' approach, which is associated with our revisionist dualist approach to neuroscience. On the other hand, our inquiry may be driven by the idea that the brain itself is an instrument with privileged access to worlds, which we have so far only minimally exploited. That is the 'microscopic' approach, which is associated with our revisionist materialist approach to neuroscience.

Of course, here we are not talking about literal telescopes and microscopes, but rather brain-oriented technologies that may function either like a telescope or like a microscope in terms of van Fraassen's distinction, depending on the epistemic relation we assume to have our brains: does the brain's normal surface functions circumscribe its capacities or does the brain possess hidden powers yet to be fathomed? Generally speaking, neuroscience approaches the brain with three types of instruments: *probes* (electrodes implanted in the cortex), *scans* (magnetic resonance imaging of brain regions) and *drugs* (targeting neurotransmitters). Brain probes may be used either as part of surgery to address a diagnosed mental disorder or in a more exploratory vein to study subjects' responses. Similarly, brain scans may be targeted to specific areas of the brain associated with a medical condition or they may provide a comprehensive survey of the brain's blood flow. With regard to drugs, the oldest

instrument available to regulate brain function, the very same drug may be seen as a cure for an existing disorder or as enhancing the performance of a normal subject. In each of these three cases, the former option represents a telescopic approach and the latter a microscopic approach to the brain. Requarth (2015) has provocatively characterised these differences in the use of neuro-technologies as, respectively, 'medicalisation' (for telescopic) and 'weaponisation' (for microscopic).

More relevant for our purposes are the radically different conceptions of the brain that underwrite these two approaches. In the telescopic case, the brain is seen as the *governor* of the body to which it is attached, with its sensory input aimed primarily at managing that body. In the microscopic case, the brain is seen as a *transducer* of external stimuli into habitable worlds, very few of which we ever realise as embodied beings. Of course, the brain is normally seen as both to some extent, but Kant played an important role in restricting the latter interpretation of the brain, which has had a profound impact on the subsequent development of neuroscience. Early in his career Kant stigmatised his older contemporary Emanuel Swedenborg, a theologically minded engineer who first scoped out the powers of the frontal lobe of the cerebral cortex, as a 'spirit-seer' for suggesting that our brains might be specially designed to know God and the supernatural more generally (Fuller 2014). Kant's encounter with Swedenborg's work turned out to be a touchstone for his later more famous work, which set limits on the claims of 'pure reason', a key moment in the institutionalisation of the modern fact-fiction distinction. Kant's rejection of Swedenborg also survives in latter-day scepticism about neuroscience's epistemic prospects, which would rate them merely in terms of telescopic inquiry.

Under the telescopic condition, neuroscience's success is judged by the extent to which our knowledge of the brain's workings makes sense of how we normally operate. Thus, one looks for regions or patterns that correspond to phenomena that we have associated with, say, 'speech' or 'memory' even before we knew anything about how the brain works. The standard of 'normal' here is 'sociological' in the broad sense of publicly observable interpersonal judgements. For example, much of what passes for 'neuromarketing' research scans the brains of people responding to marketing pitches, even though the resulting knowledge largely confirms earlier experiments confined to spoken and other behavioural responses, which in turn largely matched people's actual purchasing patterns. As for behaviours that do not conform to conventional interpersonal judgements—either with positive or negative consequences—they are presumed to correspond to some deviant brain function. However, if a subject has been deemed a 'genius', 'deficient', or 'ill' based on independent behavioural criteria, then it becomes the source of significant puzzlement if nothing out of the ordinary is

discovered about the subject's brain upon examination. That such situations do arise leads those who promote the telescopic view of neuroscience to further downgrade their epistemic expectations for the field, since a 'normal-brained' deviant suggests a preponderance of non-brain factors in determining that person's social-psychological status.

In contrast, under the microscopic condition, the standard of neuroscientific success is more ambitious: it is judged by the researcher's ability to identify areas of the brain whose stimulation enables activity of sustained novelty that the brain might not normally exhibit. The brain may be performing somewhat below a standard of which the organ is capable, even if that performance is functional in its environment. Unlike the telescopic condition, which basically envisages neuroscience as the materialist correlate of folk psychology, the microscopic approach envisages the brain as an organic technology that is semi-detached from its personal possessor, but which nevertheless contains the potential to extend its possessor's psychic powers indefinitely.

In this respect, the brain is something that we literally 'use', notwithstanding the association of that way of putting things with an 'old look' dualism. Indeed, the great twentieth century pioneers of neuroscience who directly probed the brain—Charles Sherrington and two of his notable students, Wilder Penfield and John Eccles—are canonically represented as mind-body dualists (e.g. Wickens 2015). However, they better fit with our 'new look' materialism and should be understood as making something closer to a proto-cybernetic point. When Sherrington originally depicted the brain as a very elaborate telephone exchange switchboard, an image popularised by his students, it was to suggest that as a brain matured—or a mind became autonomous—the switchboard operator was gradually internalised as part of the brain's normal operation, thereby enabling the organ to function not merely as a governor but as a transducer.

Commitment to a telescopic or a microscopic vision of neuroscience may be seen in how one interprets the idea of the brain's 'innate capacity'. Perhaps the least controversial neuroscientific appeal to the expression is in discussions of 'cognitive reserve' to refer to the mind's resistance to brain damage, reflecting both the holistic and regenerative character of the organ. While the amount and kind of cognitive reserve varies across individuals, the concept itself fits comfortably in the ambit of a telescopic approach to neuroscience. In contrast, when William James (1914) proposed that the brain had 'reserve energy', he was referring to the part of the brain that is not used at any given time. He made this claim in a popular lecture in which he exhorted the audience to make use of this reserve in order to live a more productive existence. It is clear from the context that he meant that people should put themselves in challenging situations where they are forced to think differently about the world, a prescription in line

with the 'rugged individualism' popularised by James' former student, Theodore Roosevelt, then President of the United States. To be sure, James' views about the sources of the brain's reserve energy remained quite unresolved at the time of his death. At various points, he took seriously paranormal phenomena, a hereditary unconscious as well as the idea that we normally use only 10% of our brains. Seen in retrospect a common thread in his thinking was that neuroscience should aim to enable our brains to make us more than who we have been—the hallmark of a microscopic approach to the discipline.

It is worth noting that James' imperative did not exactly go unheeded, even though this legacy has been largely erased from canonical histories of neuroscience (e.g. Wickens 2015). In particular, what one might call a 'Jamesian' attitude towards the characteristic neuroscience technologies flourished in the second half of the twentieth century, informed by the idea that the brain's underutilisation was a result of 'blockages' that inhibited the free flow of neural connections. Such thinking could be found in explicitly counter-Freudian claims that creativity in both art (Kubie 1967) and science (Maslow 1966) was not the result of sublimation—which involves a repression and channelling of base instincts—but of 'de-sublimation', so to speak. On this view, if one could break deep-seated mental habits (aka neuroses) that inhibit variation in performance, then the wellsprings of creativity could be unleashed. This move is comparable to the turn against 'representationalism' in early twentieth century art, literature and music, in which aesthetic interest was shifted from using a medium to capture a preconceived idea to exploring the medium's full expressive capacity—the relevant medium in this case being the brain (Fuller 2013).

The move is also reminiscent of what one of the original neo-liberal political economists Alexander Rüstow dubbed 'liberal interventionism', whereby the state would not simply regulate already existing markets but marketise non-market sectors of society by removing the legal protections that had allowed the formation of monopolies, which only served to stifle the free flow of goods, services—and ideas (Jackson 2009). Thus, neurological disinhibition—typically involving psychoactive drugs and/or electrical stimulation—may be seen as providing the psychodynamic analogue of what Theodore Roosevelt himself had called 'trust-busting' in the economy. Arguably this link between disinhibition and marketisation was unconsciously acknowledged in the widespread use of Jeffrey Sachs' phrase 'shock therapy' to capture the rapid neo-liberal reform of socialist economies in the late 1980s. In any case, in its heyday in the 1950s and 1960s, often with financial support and political cover from national intelligence agencies, this strand of microscopic neuroscientific inquiry was by today's standards quite adventurous, if not reckless, with subjects who underwent

various forms of extreme treatment (Winter 2011: Chap. 4; Langlitz 2012: Chap. 1).

In the late 1960s, just before the implementation of a more scrupulous ethics regime on the treatment of human—and some animal—subjects consigned much of this research to an ignominious place in the annals of neuroscience, a magnum opus was published that presaged a ‘psychocivilized society’, whereby remote control electrical stimulation of multiple brains in concert would lay the foundation for a harmonious world order (Delgado 1969). Since that time, generally speaking, development of brain control technology has been limited to the restoration of some normal human powers to the disabled and to what is increasingly known as ‘drone warfare’ (Horgan 2005). However, perhaps reflecting a bit too much success along these lines, there are now worries that we may be sleepwalking into a ‘cyborg future’, whereby the visible success of these technologies encourages ‘normal’ people to want to replace parts of their body with prosthetic extensions or offload functions of their brain to computer-based devices, or even self-identify with avatars in cyberspace (Wittes and Chong 2014). In effect, the ambitious microscopic vision of neuroscientific inquiry is returning through a normative backdoor as people rethink ‘ability’ in terms of what they would like to be and do rather than what is natural for them to be and do.

Before concluding with a more futuristic look at the prospects for neuroscience as an organic technology, let us consider an analogue to the tele-/micro-scope distinction in more conventional silicon-based information technology. To be sure, the ontological baseline of the brain and the computer are radically different. However, one characterises the ‘information’ coded in the brain, it occurs in neural networks of varying densities, whereas the information coded in a computer occurs in data sets that are accessible by any number of algorithms. The advent of ‘big data’ has highlighted just how much our orientation to the computer as a knowledge producer has been more telescopic than microscopic.

For the past quarter-century the field of ‘knowledge management’ has flourished in business schools by promoting the image of data as something that should be ‘mined’ (Fuller 2002: Chap. 1). An important implication of this image is that some data will be retained as ‘mineral’ but most will be discarded as ‘ore’. Algorithms are designed to survey big data streams in search of specific patterns that the end-user has already identified of relevance. But as these algorithms ‘cut through the noise’, the end-user never really sees the full range of data available—including those that might be of interest. This point has taken on an added significance with the advent of personalised health-oriented and performance-based ‘self-tracking’ devices. Here one might say a kind of ‘telescopic fallacy’ is in full force: it is like trying to cure myopia by creating lenses that enable you to see more clearly what you can already see indistinctly but

not what you have never seen at all. Thus, in the name of improving accuracy, bias is effectively reinforced.

But there is also a microscopic orientation to big data. It involves ‘data surfacing’ rather than ‘data mining’. This contrast is a coinage of the leading Silicon Valley cybersecurity firm, Palantir (<http://www.palantir.com>), which structures large amounts of data in ways that enable the end-user to search for patterns inductively—that is, with a relatively loose initial sense of what might be of interest. Whereas data-mining reinforces your cognitive biases, data-surfacing aims to extend your cognitive horizons. A characteristic feature of data-surfacing techniques is that they require significant human input through the life of the algorithm in order for the output to be represented in a way that enables the end-users to gain maximum advantage. Thus, Palantir prides itself in selling not only data-surfacing platforms but also installing its own engineering staff to enable end-users to discover things about the data at their disposal that they perhaps would not have thought of looking for before. While these services are compelling in the concrete context of anticipating the next terrorist attack, they underscore at a conceptual level how human intelligence may be enhanced—as opposed to being merely replaced or disciplined—by machine intelligence.

3 Conclusion: does the future of neuroscience lie in cognitive agriculture?

Opposing claims to *efficiency* have been made on behalf of the brain vis-à-vis the computer as (for want of a better expression) a ‘knowledge producer’. When the stress is placed on processing speed and average accuracy of outcomes in a specified domain, the computer looks more efficient. But when the stress is placed on energy usage and average accuracy of outcomes across many domains, the brain looks more efficient. The two sets of criteria are easily converted into alternative standards for evaluating ‘intelligence’ in agents more generally (cf. Hernandez-Orallo 2017). In that respect, conflicting popular judgements about the relative intelligence of humans and computers trade on differing intuitions of what intelligence is. Moreover, as I suggested in the first section, these intuitions themselves may represent alternative paradigms for organising energy efficiently.

Nevertheless, brains and computers share a key feature that offers hope to those interested in computational models of the brain: a trade-off between processing speed and energy usage in terms of efficiency. Both require more energy to process the same information more quickly (Hruska 2014). And in terms of baseline energy needs, the brain is likely to remain a much more efficient knowledge producer than a

computer for the foreseeable future. A sense of the challenge facing silicon advocates is that recently a supercomputer programmed with a neural network required 40 min to simulate 1 s of processing in a brain 2% of the size of a normal human brain (Whitwam 2013). This suggests that while it may make sense to develop supercomputers capable of surpassing human performance in a range of specific tasks, it would be an ecological disaster to try to create an artificial intelligence capable of approximating the performance of the entire human brain. The prospect brings to mind the denouement of the 2014 film *Transcendence*, in which the first human to have his brain uploaded into a supercomputer manages to short-circuit the entire planet.

One way around this ecological impasse might be to upload a vast number of brain emulations into a single supercomputer, resulting in a Star Trek Borg-style hive mind. Of course, we have yet to computationally emulate an entire human brain, let alone upload it into a supercomputer to acquire a second life. But this has not stopped people from thinking about how such a development might alter the attribution of rights and responsibilities under the law (Fuller 2016b), not to mention human evolution more generally (Hanson 2016). Nevertheless, given the enormous ecological challenge faced by silicon-based computers, might it not be more sensible to return to carbon and pursue a strategy of what might be called *cognitive agriculture*? For example, one might use stem cells to grow brains or brain-like entities from dense clusters of neurons, resulting in multipurpose, energy-efficient organic knowledge producers. A harbinger may be Philip K. Dick's 1956 short story, 'The Minority Report', which was popularised by Steven Spielberg in a 2002 film starring Tom Cruise. Here we find hydroponically cultivated 'precogs', mutant offspring of drug addicts, who are repurposed to anticipate crimes because of their ability to project visions of the future based on processing multiple data streams much more efficiently than any silicon-based computer. Thus, in the spirit of Putnam's 'brains in a vat' thought experiment raised earlier in this article, the Cartesian demon is turned to some socially acceptable cognitive advantage.

Of course, Dick wrote before embryonic stem cell technology became available. In the not too distant future it may be possible to acquire a significant portion of the brain's computational power without bringing an entire human being to fruition. Whether from a legal standpoint the resulting entity would constitute a person in its own right or merely the possession of its owner is an open question. Moreover, we should expect in the meanwhile the same sort of 'morally principled' objections to the artificial growing of brains as we currently see to stem cell research more generally and genetically modified organisms. Yet, at the same time, exploratory research is underway which have led to impressive proofs of concept that DNA may be a more

efficient way of storing digital information—and not only genetic information—than the brain or even silicon chips (Rutherford 2013). The major obstacles at the moment pertain to the speed of encoding and the ease of accessing the encoded information. But were research in this area to be incentivised, cultivated strands of the 'genetic code' could become an important if not the primary means by which information in general is stored and retrieved. This 'smart organ' may be located either in one's own body or in a dedicated device. But in either case, it would constitute a parallel development that suggests a bright future for cognitive agriculture.

I do not underestimate the moral and technical objections to a political economy driven by cognitive agriculture. But in the end, an 'ecomodernist' argument—that is, one based on an innovation-driven strategy of energy conservation—is likely to prevail in the trajectory's favour, if we are determined to promote the fruits of 'advanced human civilization' into the indefinite future on planet Earth (cf. Fuller 2016c).

Acknowledgements Fuller would like to acknowledge support of the Russian Science Foundation, project number 14-18-02227, 'Social Philosophy of Science', with which he is associated as a Research Fellow in the Russian Academy of Sciences, Institute of Philosophy.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Bennett M, Hacker P (2013) History of cognitive neuroscience. Wiley, Chichester
- Boden M (2006) mind as machine: a history of cognitive science. Oxford University Press, Oxford
- Brey P (2000) Theories of technology as extension of human faculties. In: Mitcham C (ed) Metaphysics, epistemology and technology, vol 19. Elsevier, London
- Cabrera L (2015) Rethinking human enhancement: social enhancement and emergent technologies. Palgrave Macmillan, London
- Church G, Regis E (2012) Regenesi. Basic Books, New York
- Clark A (2008) Supersizing the mind: embodiment, action and cognitive extension. Oxford University Press, Oxford
- Delgado J (1969) Physical control of the mind: towards a psychocivilized society. Harper & Row, New York
- Dupuy J-P (2000) The origins of cognitive science. Princeton University Press, Princeton
- Fuller S (2002) Knowledge management foundations. Butterworth-Heinemann, Woburn
- Fuller S (2006) The new sociological imagination. Sage, London
- Fuller S (2013) The dawn of critical neuroscience. Hist Hum Sci 26(3):107–115
- Fuller S (2014) Neuroscience, neurohistory and history of science: a tale of two images of the brain. Isis 105:100–109

- Fuller S (2016a) Organizing the organism: a re-casting of the bio-social interface for our times. *Sociol Rev Monogr* 64(1):134–150 (**Special issue on ‘Biosocial Matters’**)
- Fuller S (2016b) Morphological freedom and the question of responsibility and representation in transhumanism. *Confero* 4(2):33–45
- Fuller S (2016c) Humanity’s lift-off into space: prolegomena to a cosmic transhumanism. In: Armstrong R (ed) *Star ark: a living, self-sustaining spaceship*. Springer, Berlin, pp 383–393
- Fuller S, Lipinska V (2014) The proactionary imperative: a foundation for transhumanism. Palgrave Macmillan, London
- Gardner H (1985) *The mind’s new science*. Basic Books, New York
- Hanson R (2016) *The age of em: work, love and life when robots rule the earth*. Oxford University Press, Oxford
- Hernandez-Orallo J (2017) *The measure of all minds: evaluating natural and artificial intelligence*. Springer, Berlin
- Horgan J (2005) The forgotten era of brain chips. *Sci Am* 293(4):66–73
- Hruska J (2014) The human brain’s remarkably low power consumption, and how computers might mimic its efficiency’ *ExtremeTech* (9 July). <https://www.extremetech.com/extreme/185984-the-human-brains-remarkably-low-power-consumption-and-how-computers-might-mimic-its-efficiency>. Accessed 9 Feb 2018
- Jackson B (2009) At the origins of neo-liberalism: the free economy and the strong state, 1930–47”. *Hist J* 53:129–51
- James W (1914) *The energies of men*. Moffat, Yard and Company, New York (**Orig. 1906**)
- Kubie L (1967) *Neurotic distortion of the creative process*. Farrar, Straus and Giroux, New York
- Kuhn T (1962) *The structure of scientific revolutions*. University of Chicago Press, Chicago
- Langlitz N (2012) *Neuropsychedelica: the revival of hallucinogen research since the decade of the brain*. University of California Press, Berkeley
- Malapi-Nelson A (2017) *The nature of the machine and the collapse of cybernetics*. Palgrave Macmillan, New York
- Maslow A (1966) *The psychology of science*. Harper and Row, New York
- Mason LR (2015) *Rewiring the brain to create new senses*. Motherboard. https://motherboard.vice.com/en_us/article/8qxjmg/rewiring-the-brain-to-create-new-senses. Accessed 9 Feb 2018
- Pickering A (2010) *The cybernetic brain*. University of Chicago Press, Chicago
- Pinker S (2002) *The blank slate*. Random House, New York
- Putnam H (1982) *Reason, truth and history*. Cambridge University Press, Cambridge UK
- Requarth T (2015) This is your brain. This is your brain as a weapon. *Foreign Policy*. <http://foreignpolicy.com/2015/09/14/this-is-your-brain-this-is-your-brain-as-a-weapon-darpa-dual-use-neuroscience/>. Accessed 9 Feb 2018
- Rubin B (2009) Changing brains: the emergence of the field of adult neurogenesis. *BioSocieties* 4:407–424
- Rutherford A (2013) *Creation: the origin of life/the future of life*. Penguin, London
- Smail D (2008) *On deep history and the brain*. University of California Press, Berkeley
- Van Fraassen B (1981) *The scientific image*. Oxford University Press, Oxford
- Weir T (ed) (2012) *Monism: science, philosophy, religion, and the history of a worldview*. Palgrave Macmillan, London
- Whitwam R (2013) Simulating 1 second of human brain activity takes 82,944 processors. *ExtremeTech* (5 August). <http://www.extremetech.com/extreme/163051-simulating-1-second-of-human-brain-activity-takes-82944-processors>. Accessed 9 Feb 2018
- Wickens A (2015) *A history of the brain from stone age to neuroscience*. Psychology Press, London
- Williams RJP, Frausto da Silva JJR (2007) Evolution revisited by inorganic chemists. In: Barrow JD (eds) *Fitness of the cosmos for life*. Cambridge University Press, Cambridge, pp 456–489
- Winter A (2011) *Memory: fragments of a modern history*. University of Chicago Press, Chicago
- Wittes B, Chong J (2014) *Our cyborg future: law and policy implications*. Brookings Institution, Washington DC